Effective ways to use statistical methods to accelerate bioassay development.

David Lansky, Ph.D.
Lansky Consulting, LLC
257 S. Union St.
Burlington, VT 05401
801 865 0155
802 861 2365 FAX
www.lanskyconsulting.com
david@lanskyconsulting.com

Outline

- Bioassay introduction
- Assay optimization precision/cost/time, delay until product shows clear promise, BUT
- Bioassay is an important guide to early product development, high precision early is valuable (low precision precludes use to guide product development w/out many replicates)
- Statistical tools for better assays:
  - Randomization
  - Experimental units
  - Blocking
  - Factorial, fractional factorial, and response surface designs
  - Confidence intervals (particularly for no-difference experiments)
  - Mixed models

What is a bioassay?

- Measurement system based on comparing responses of groups of living organisms
- Not calibration based – relative potency
- Typically noisy
- Typically laborious
- Usually want several checks on assay
  - Reference (standard) looks typical
  - Reference and test give “similar” dose-response curves
- USP prescribes additional checks
- Viral bioassays are “double trouble”

Importance of bioassay

- Close to the clinical outcome
- Often 1st to detect problems
  - stability
  - impurities
- “Well characterized product” requires a lot of experience
- Bottleneck: Often doesn’t get enough attention early enough
  - development & validation
  - carpel tunnel by a key analyst happens
  - bioassay analysis software

Bioassay: Key idea – relative potency

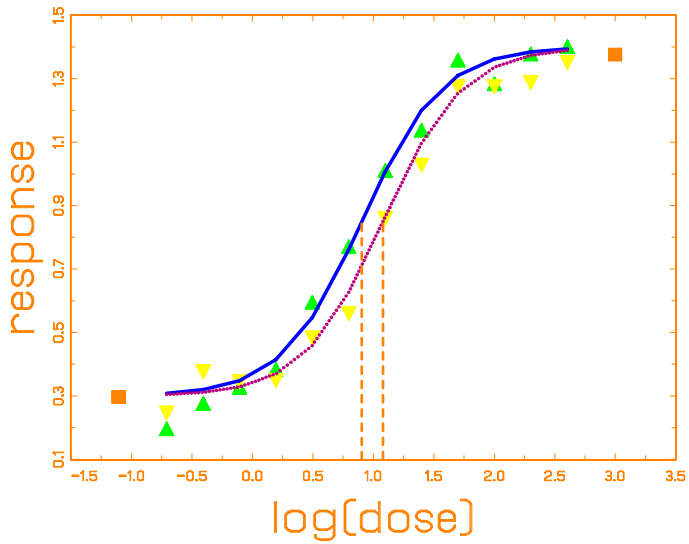
Standard and test samples close together

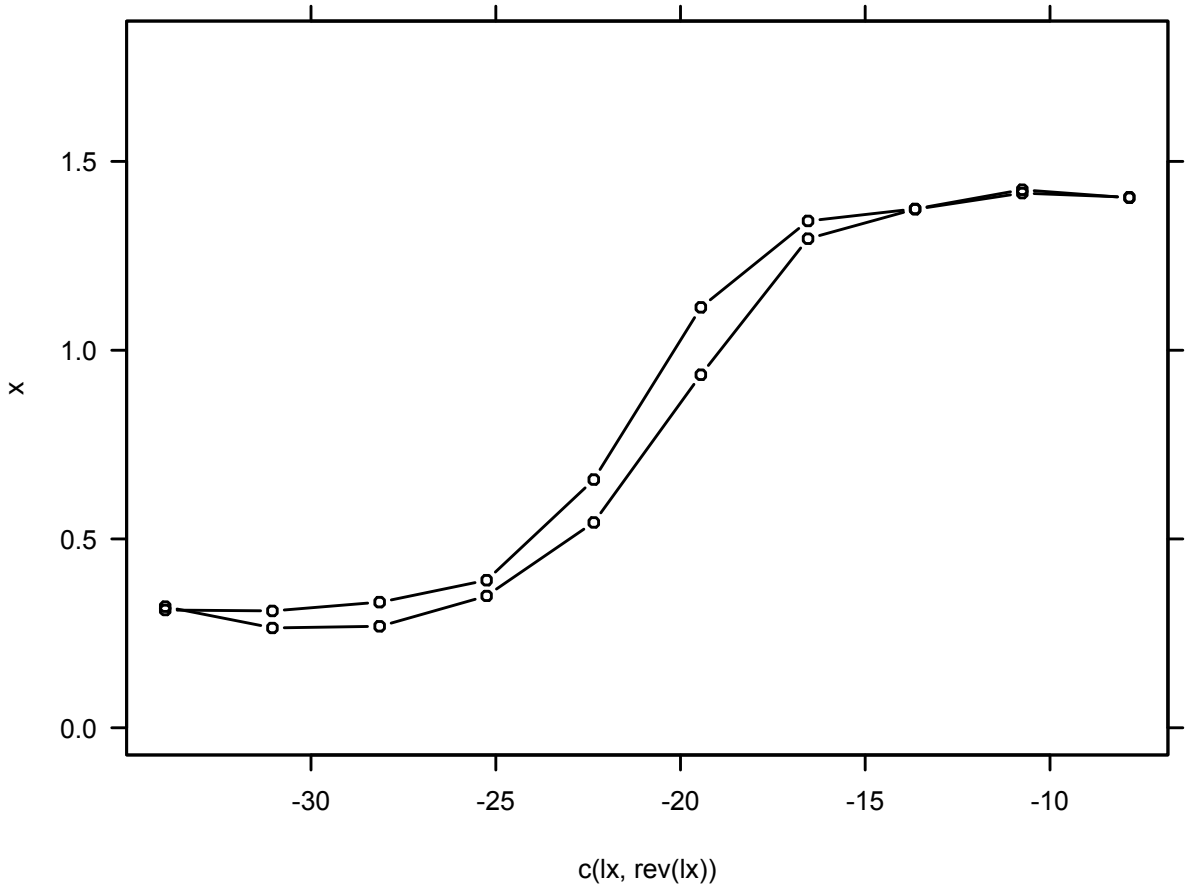
If the curves have the same shape

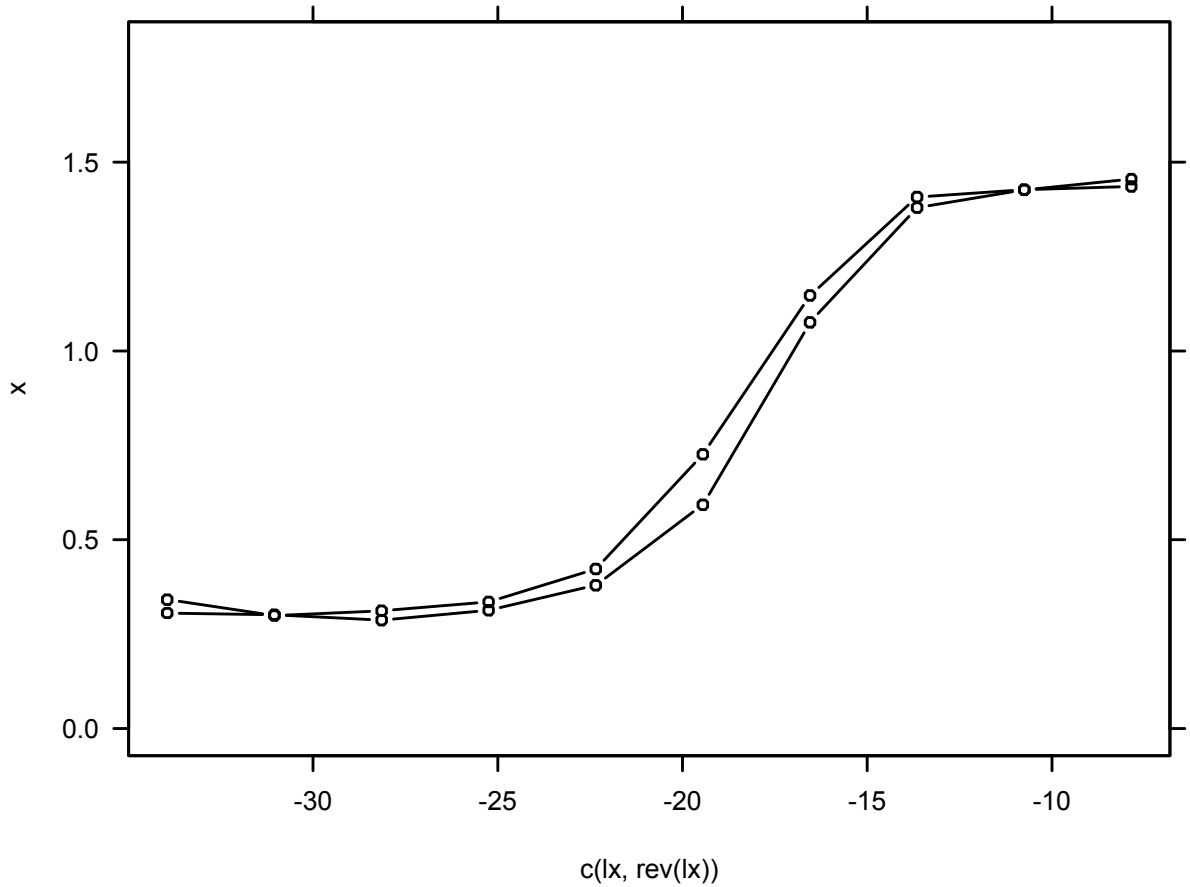
- no evidence that compounds differ
- no evidence that organisms are not comparable
- impose common slope and asymptotes
- estimate horizontal displacement and interpret as relative potency

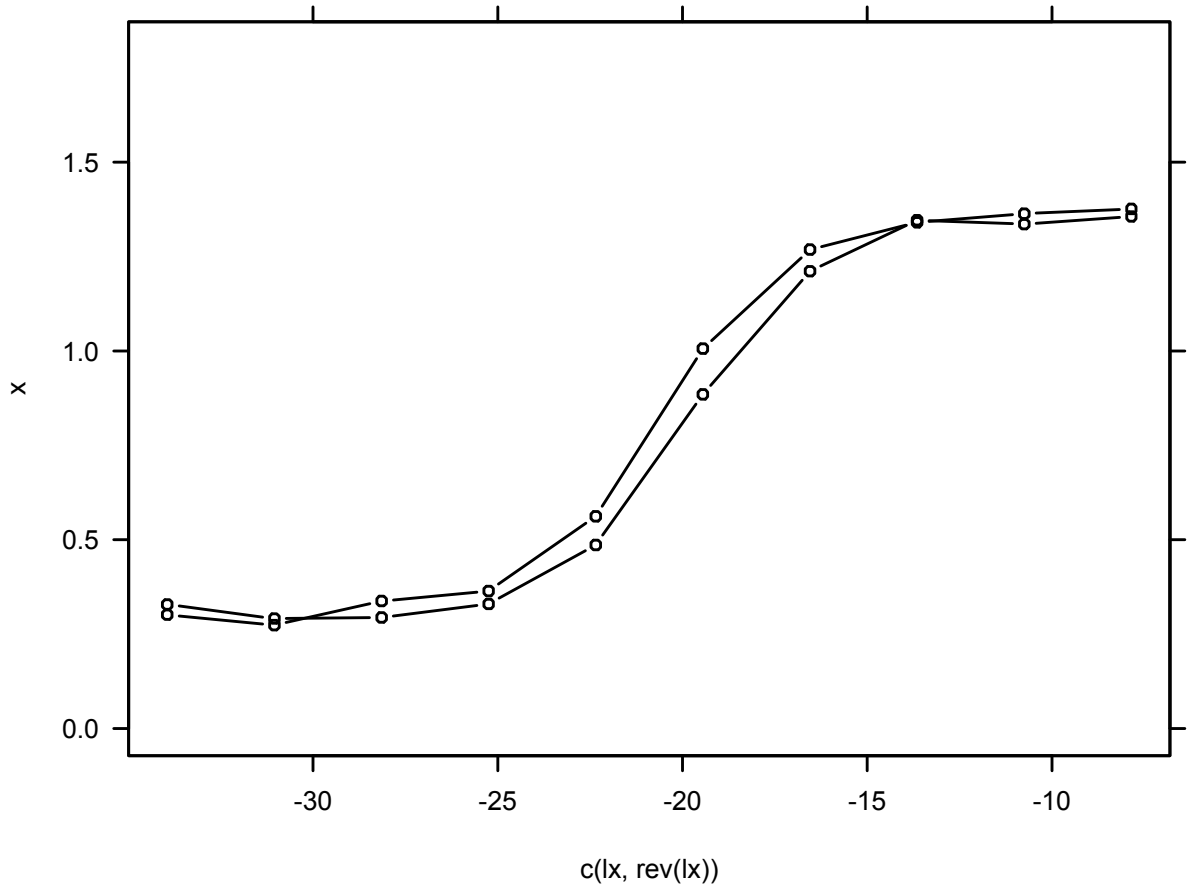
Removes variation in curve shape from assay to assay

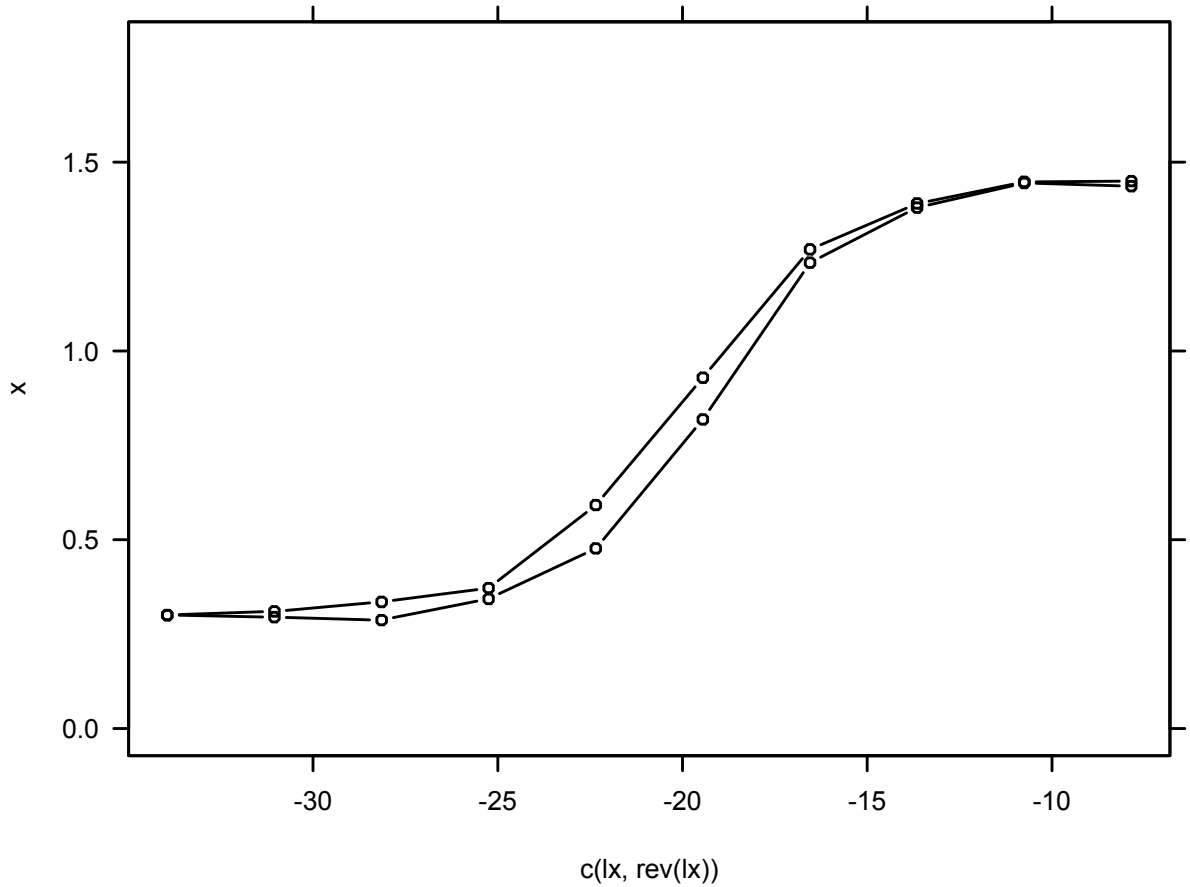
What's constant over the next 15 slides?

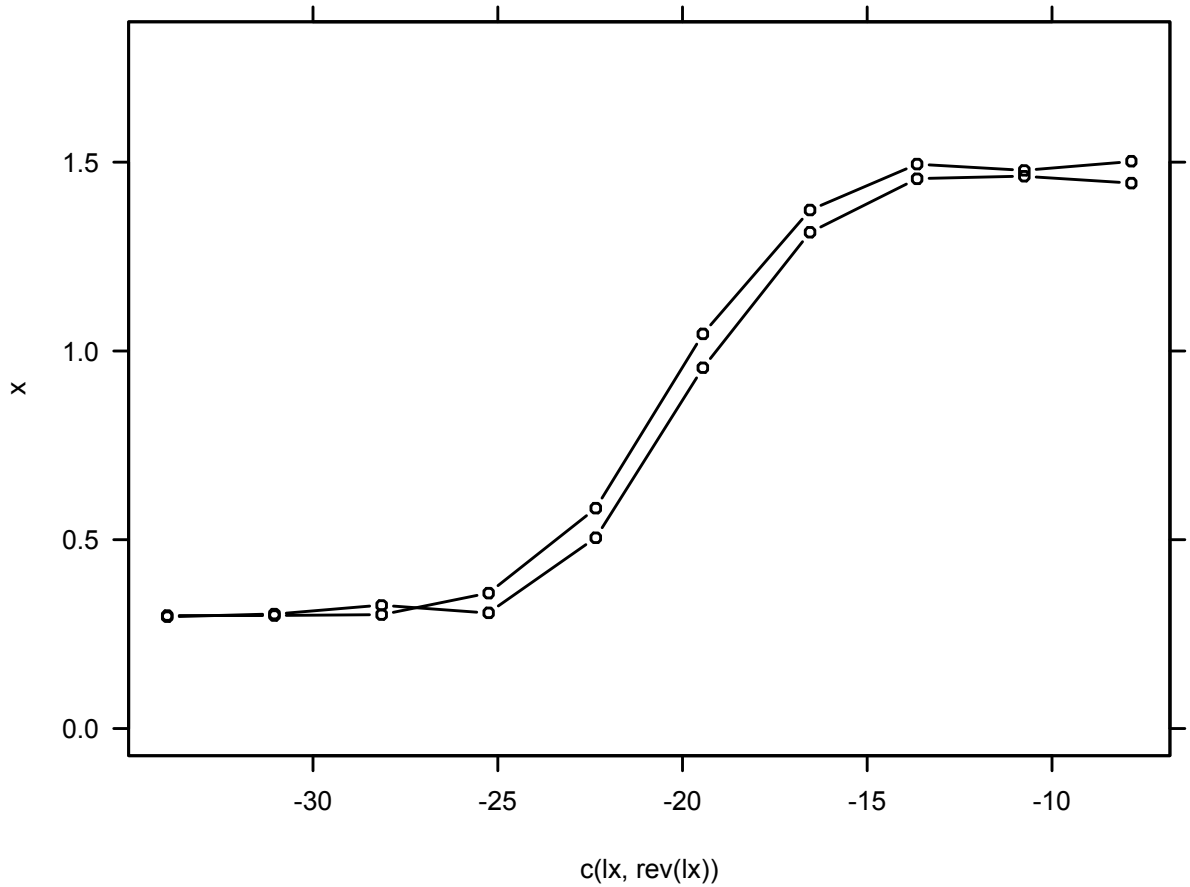


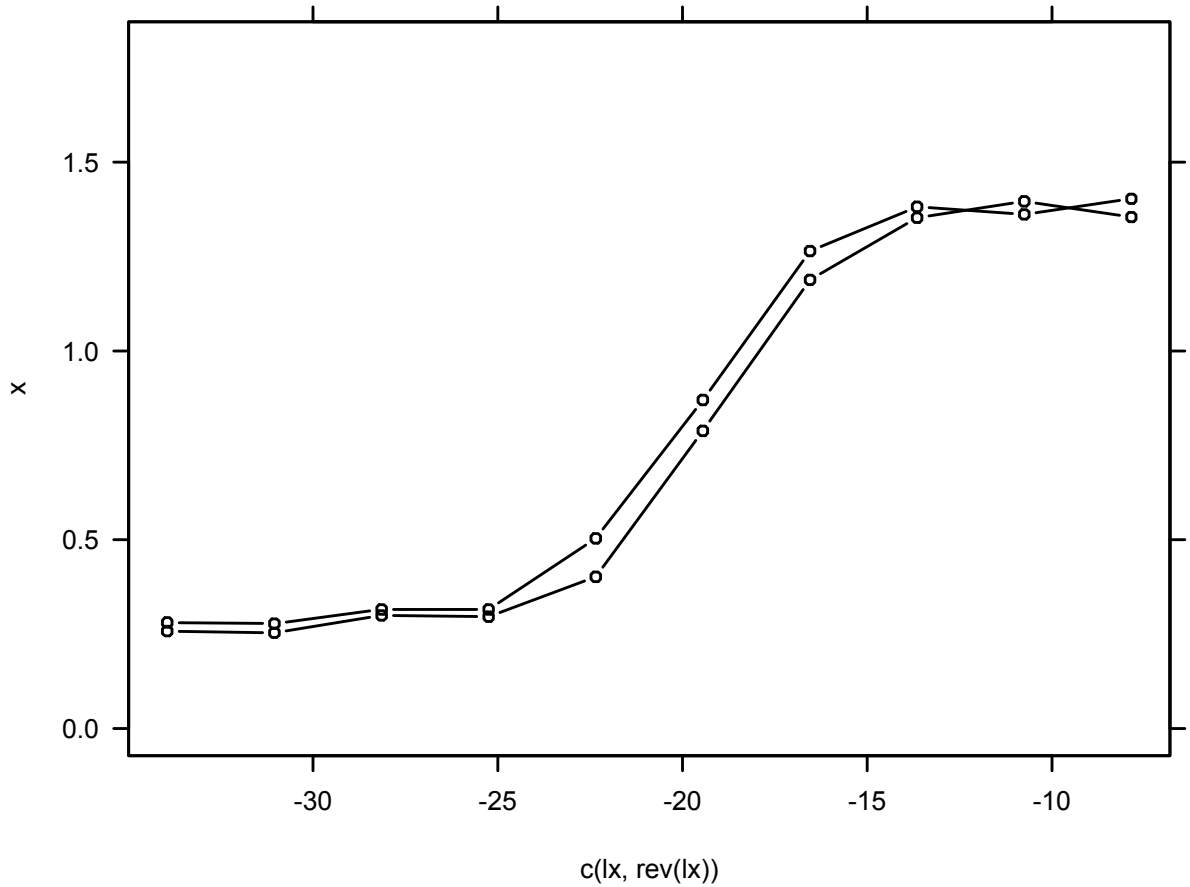


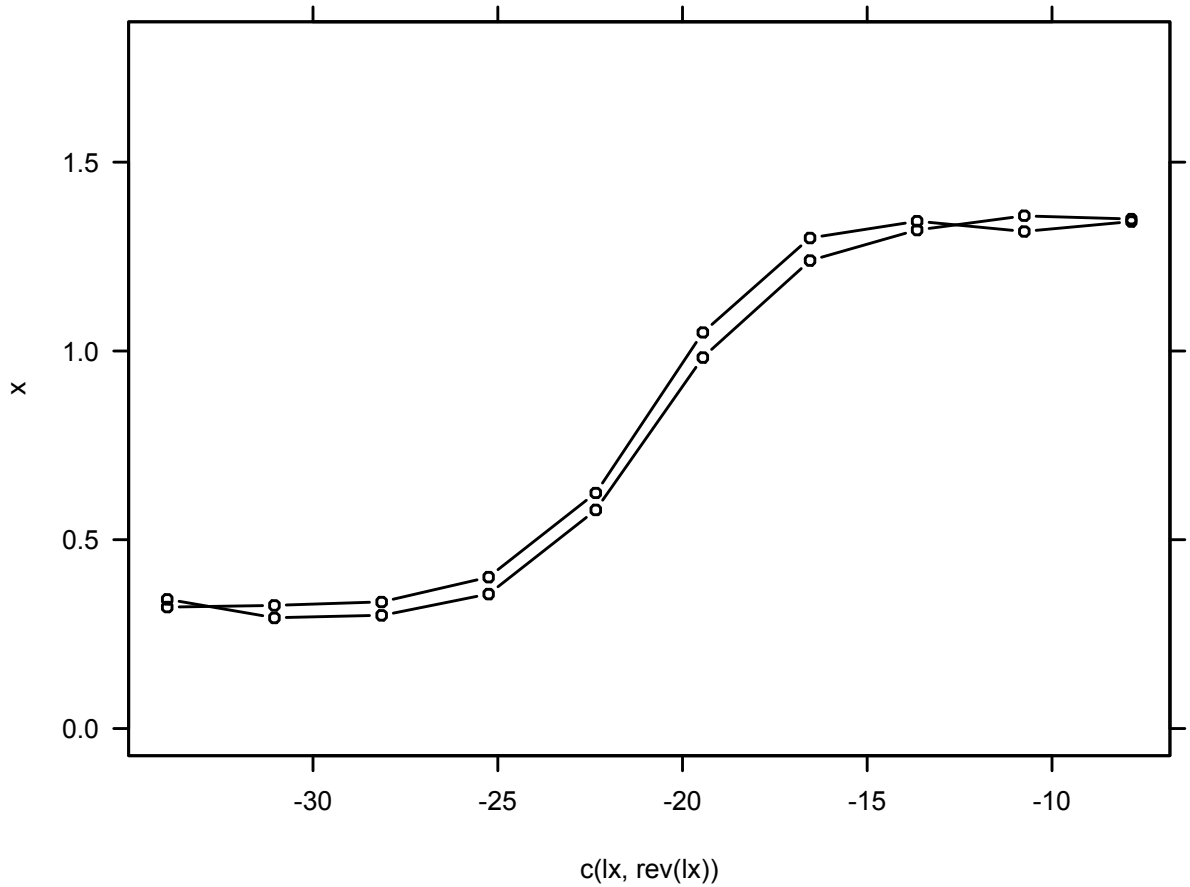


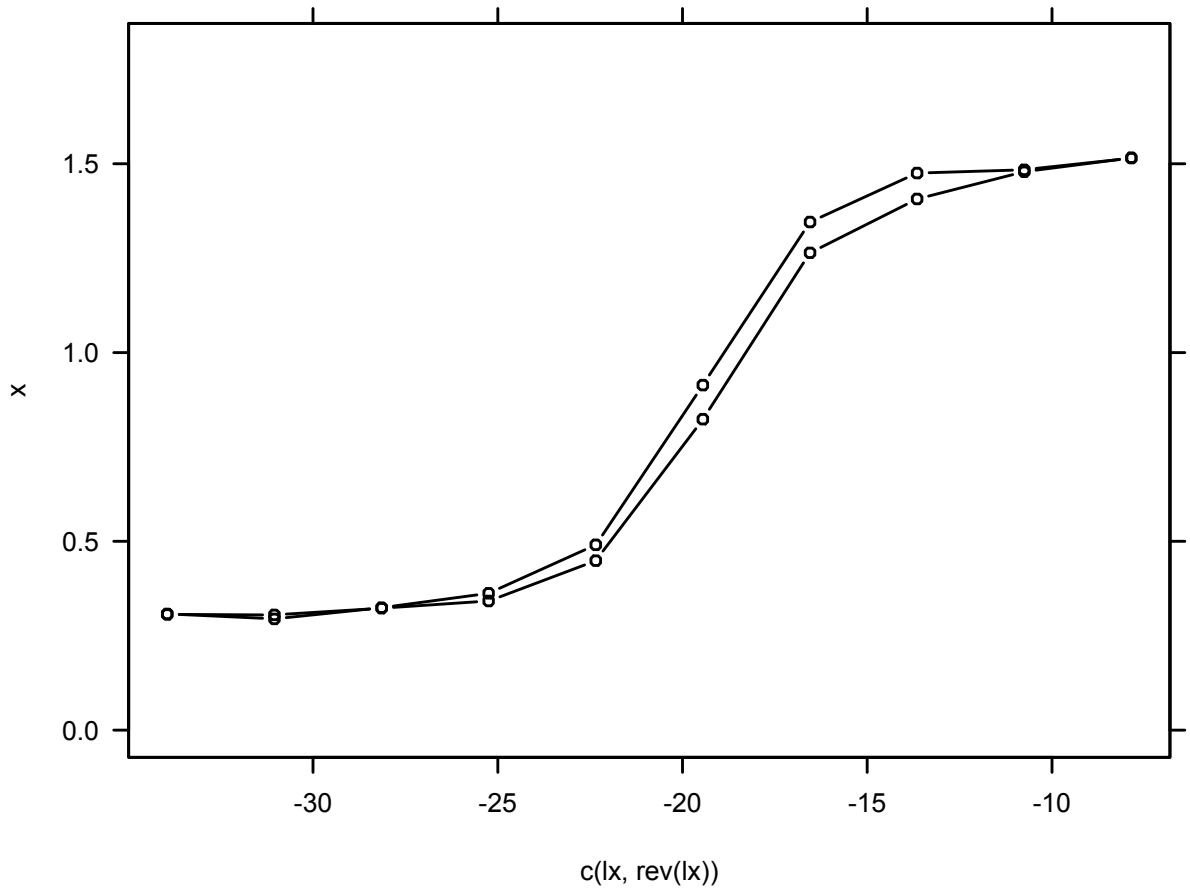


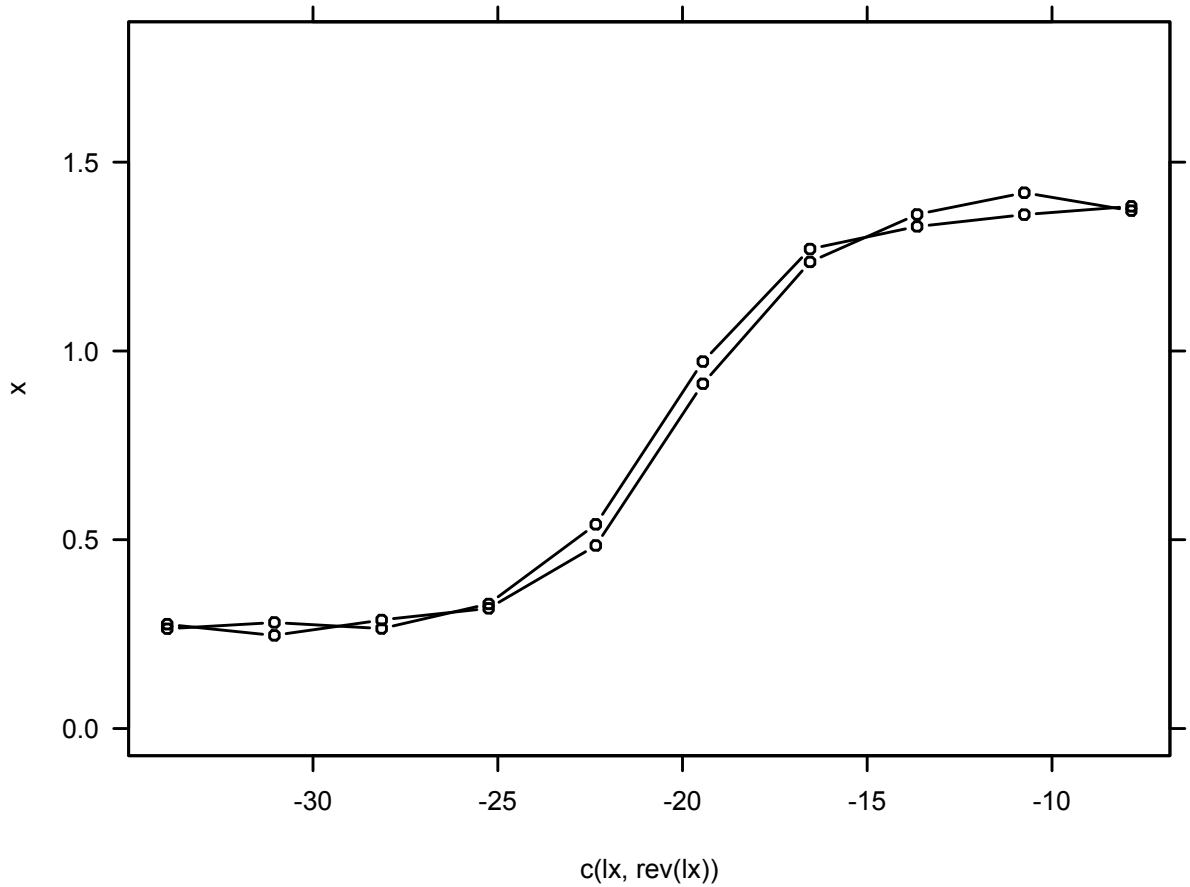


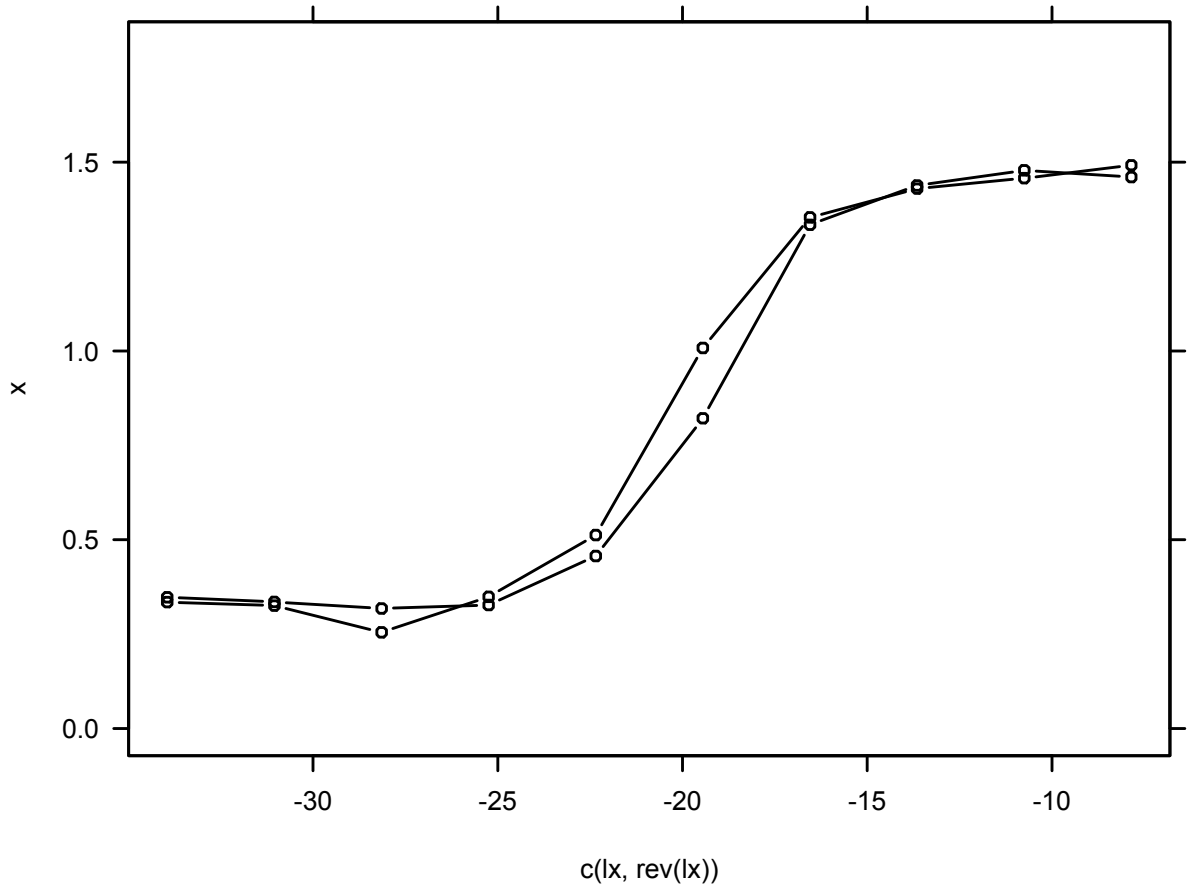


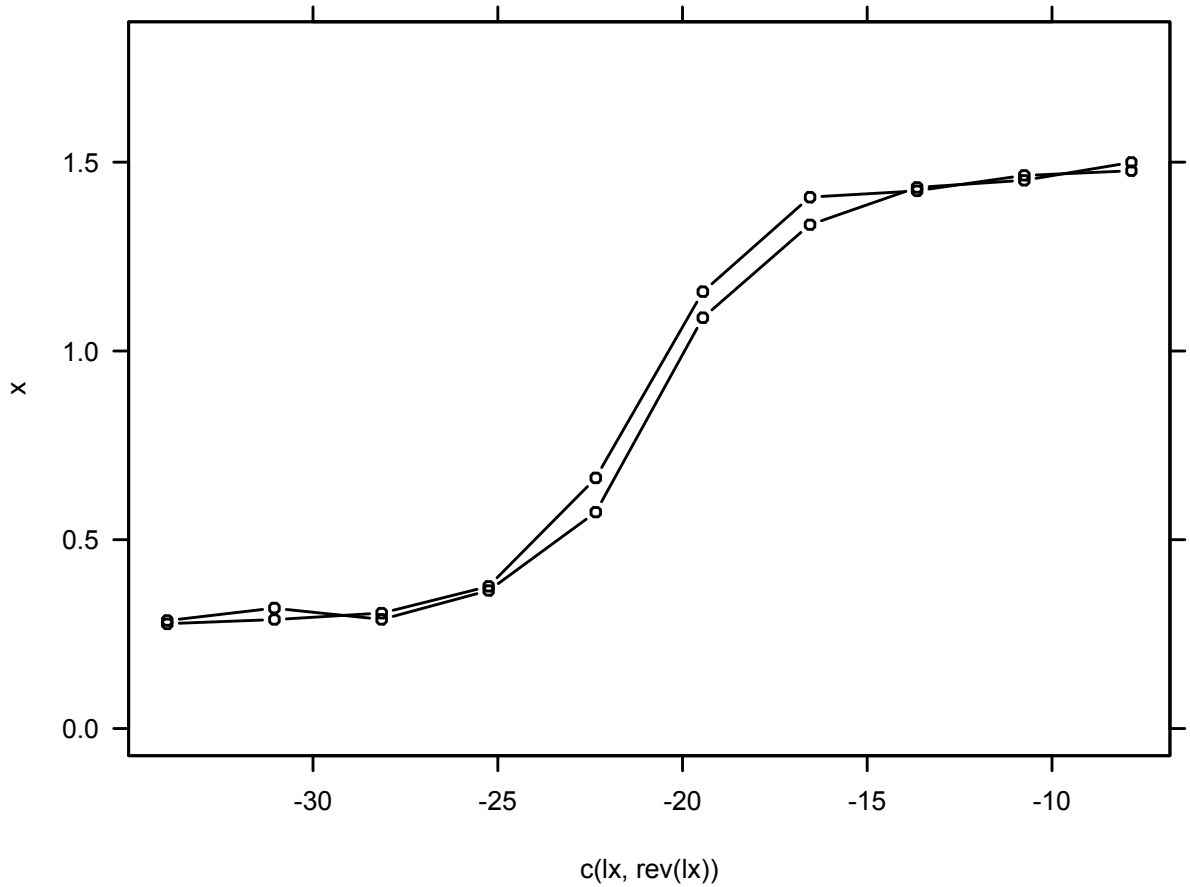


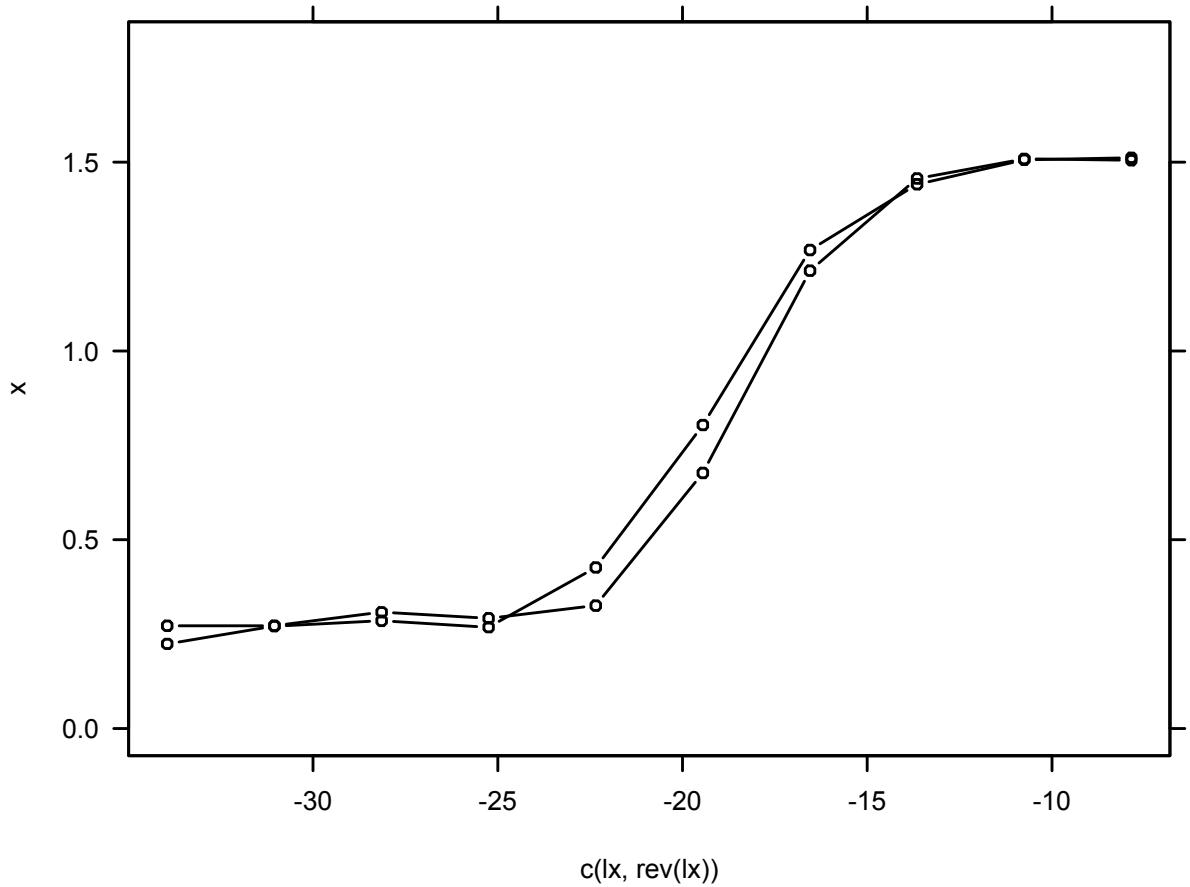


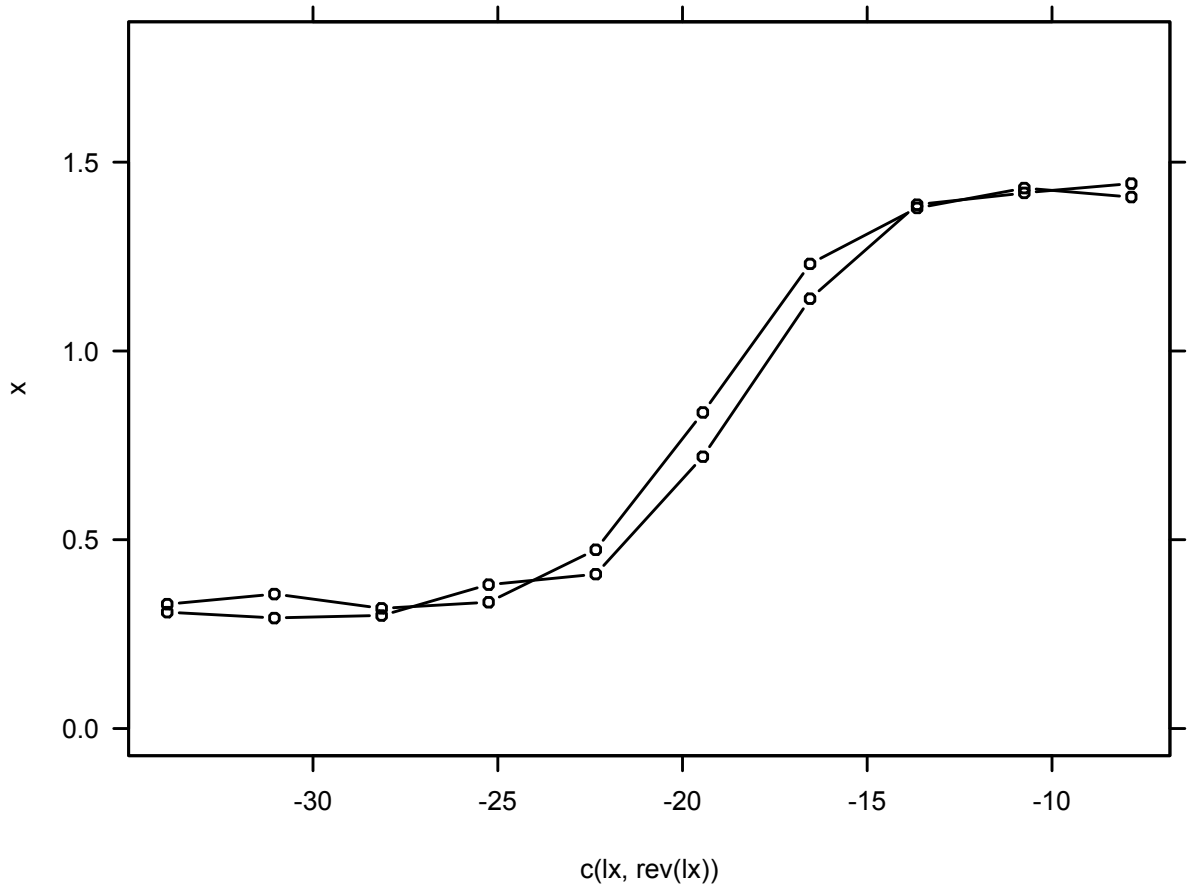


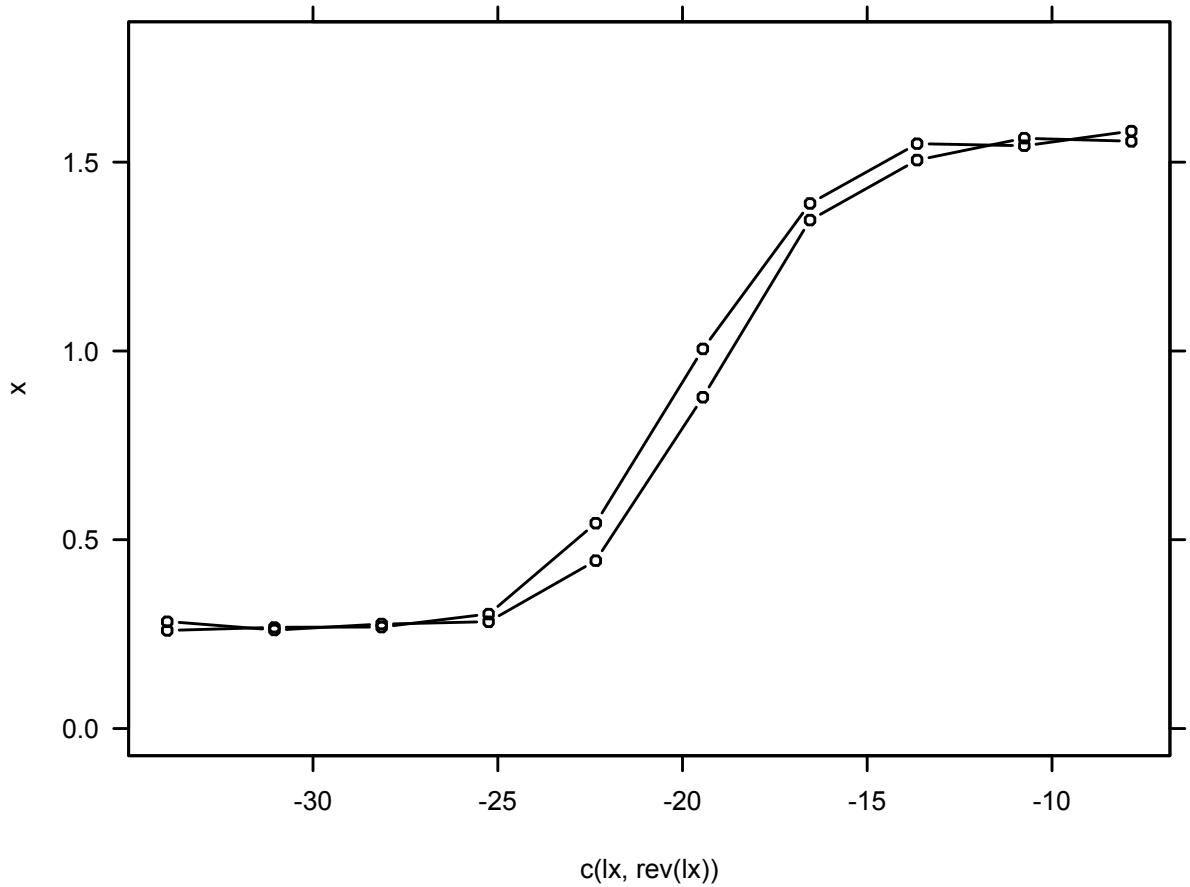


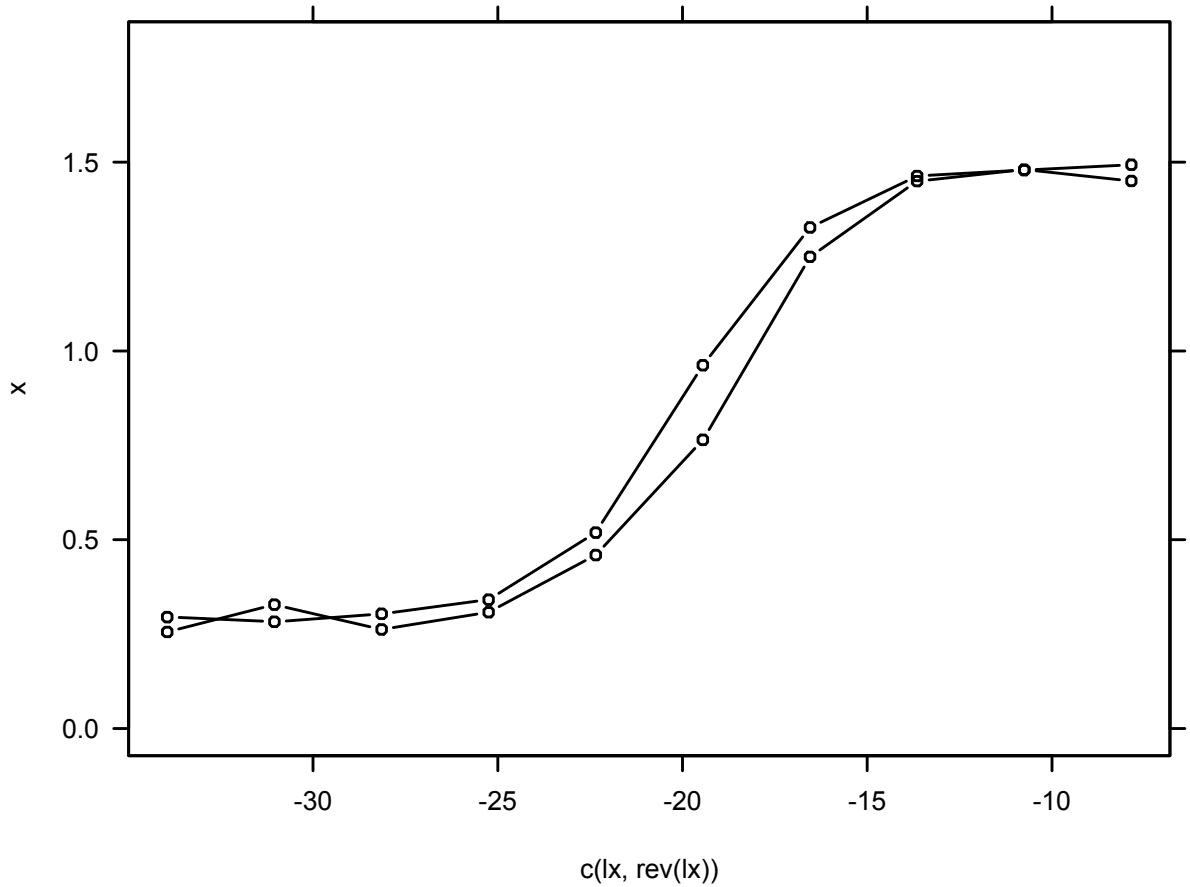












What precision is needed on potency when?

- Regulatory: must have validated bioassay by Phase III (typically need an RSD of 7% or 15% for protein or vaccine)
- Many bioassays have poor precision until late Phase II
- When evaluating changes in product production process or assay it is hard to find changes in potency
- To have an 80% chance of detecting a process improvement of **δ** using a 5% test for an assay with a Relative Standard Deviation (RSD) on potency of 50% will require **n** assays

Conclusions:

1) Small studies will consistently fail to find modest but important process improvements

2) "No difference" studies will almost always succeed

N	σ^2/δ^2	δ
7	1	50%
25	4	25%
155	25	10%

Uses of Bioassay pre-Phase III

- Dosing animals in pre-clinical studies
  - toxicity, etc.
  - efficacy
- Refining production process
- Product formulation
- Preliminary product stability
- Product uniformity
  - between lots
  - within lots

Fast Bioassay Development

- Get organized about the process
- Use statistical tools for development
  - randomization
  - good design of the assay (i.e.; plate layout)
  - good analysis (exploit blocks, mixed models)
  - factorial experiments to find important factors
  - response surface to optimize critical factors
  - study precision of the system early
- For production use of the assay: recast approach to assessing similarity

Basic Statistics: Randomization

- Why should we randomize?
  -
  -
  -
- Why don't we randomize
  -
  -
- What does the FDA do with non-randomized clinical trial results? Should this standard apply to assays?

Basic Statistics: Experimental Unit

- The experimental unit is the smallest unit randomly allocated to a distinct treatment.
- Examples:
  - Mice: dose/sample assigned to cage
  - Mice: sample to mouse, dose to cage (split-plot)
  - 96 well plate
    - samples assigned to?
    - dose assigned to?

Basic Statistics: Expt. unit for sample?

	A1	A1	A1	B1	B1	B1	R1	R1	R1	+	
	A2	A2	A2	B2	B2	B2	R2	R2	R2	+	
	A3	A3	A3	B3	B3	B3	R3	R3	R3	+	
	A4	A4	A4	B4	B4	B4	R4	R4	R4	-	
	A5	A5	A5	B5	B5	B5	R5	R5	R5	-	
	A6	A6	A6	B6	B6	B6	R6	R6	R6	-	

How to improve?

Basic Statistics: Expt. unit for sample?

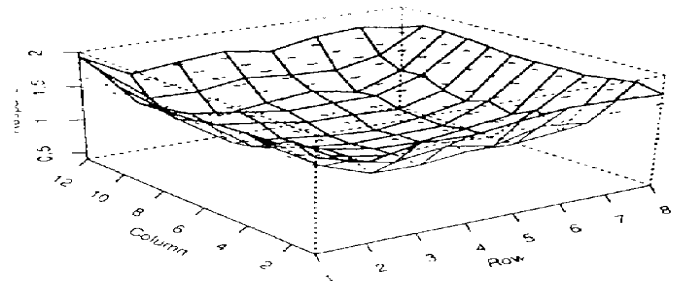
	A1	B1	A1	R1	B1	A1	R1	B1	R1	+	
	A2	B2	A2	R2	B2	A2	R2	B2	R2	+	
	A3	B3	A3	R3	B3	A3	R3	B3	R3	+	
	A4	B4	A4	R4	B4	A4	R4	B4	R4	-	
	A5	B5	A5	R5	B5	A5	R5	B5	R5	-	
	A6	B6	A6	R6	B6	A6	R6	B6	R6	-	

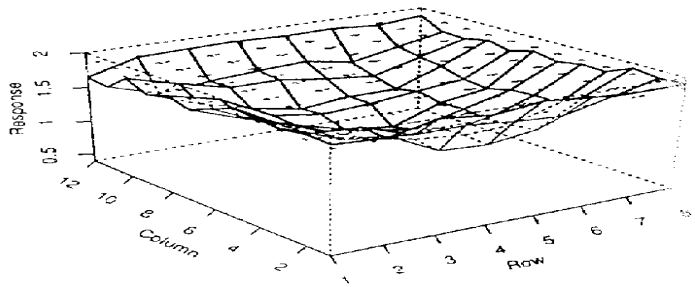
How to improve?

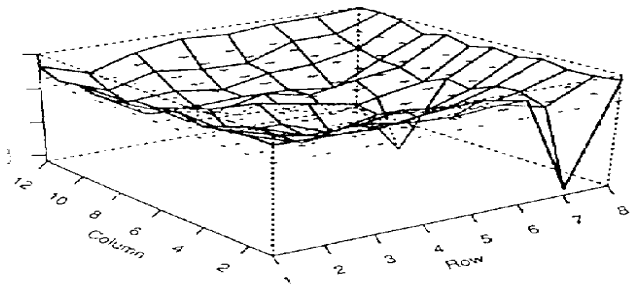
Why are these designs in use?

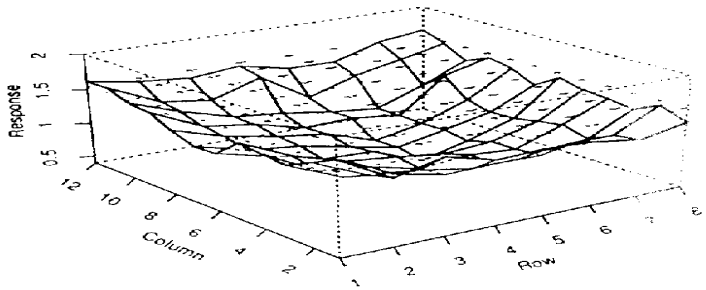
- Physical constraints of lab equipment
  - Multichannel pipettes
  - Real need to work quickly in routine ways
- Strong desire to keep lab methods consistent
  - randomization is an invitation for procedural mistakes
- Few statisticians, bench scientists, or assay managers are comfortable with experimental units
- Many bench scientists and assay managers don't ask for (agricultural) statistical input early enough
- Not enough statistical input on the design of software for lab robots

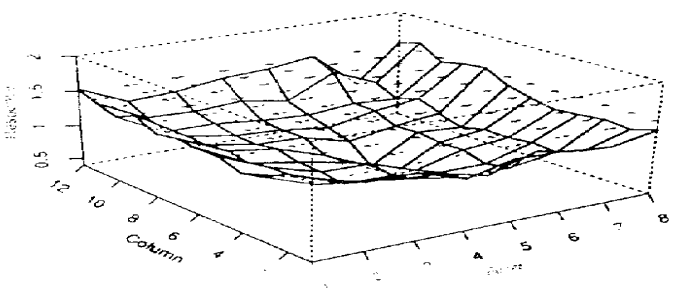
Uniformity trial #1

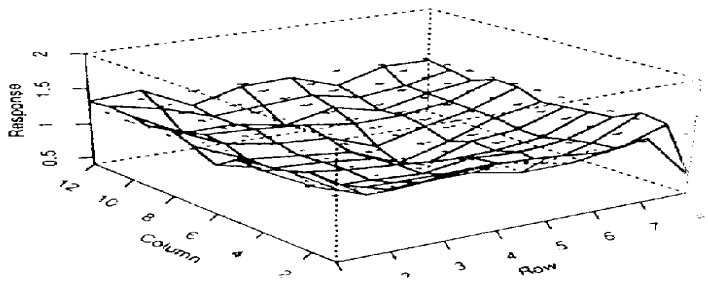












Basic Statistics: Expt. unit for sample?

	A1	R1	B1	R1	B1	A1	B1	R1	A1	+	
	A2	R2	B2	R2	B2	A2	B2	R2	A2	+	
	A3	R3	B3	R3	B3	A3	B3	R3	A3	+	
	A4	R4	B4	R4	B4	A4	B4	R4	A4	-	
	A5	R5	B5	R5	B5	A5	B5	R5	A5	-	
	A6	R6	B6	R6	B6	A6	B6	R6	A6	-	

What are these groups of columns?

Basic Statistics: Blocks

- Blocks are EVERYWHERE
- Exploiting blocks is a powerful design technique; it is **THE** core idea in bioassay analysis
  - Variation among blocks is removed from the analysis
  - Comparisons within blocks are much more precise
- Goal: associate unwanted (uncontrollable) variation with blocks
- We are (almost) never interested in individual block means, only in the comparisons within block
- Think of variation among blocks as random, eg:
  - plates, cages, days

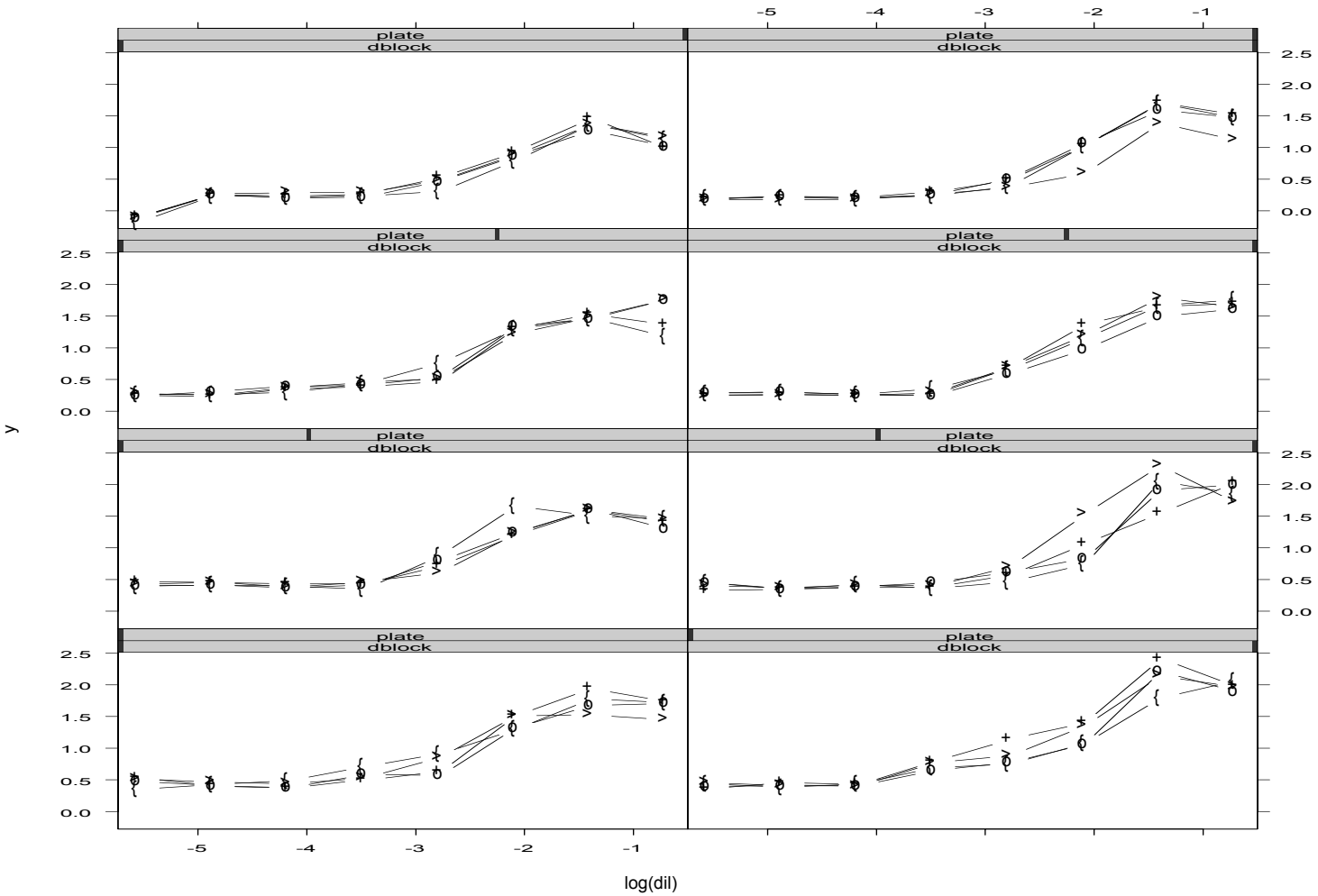
A split-plot with samples randomized to row and dose randomized to well would be:

	A3	A6	A10	A2	A5	A1	A4	A9	A8	A7	
	R6	R9	R1	R2	R5	R10	R3	R8	R7	R4	
	B1	B5	B2	B8	B3	B10	B7	B9	B4	B6	
	R5	R1	R9	R4	R2	R8	R3	R10	R6	R3	
	A4	A2	A8	A7	A1	A9	A3	A10	A5	A6	
	B10	B3	B7	B2	B4	B5	B9	B6	B1	B8	

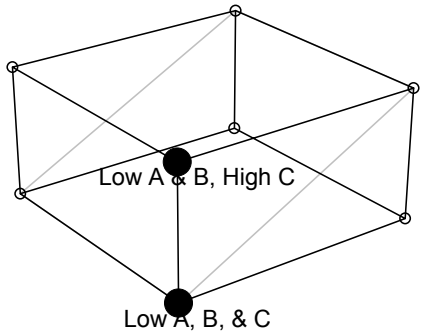
A strip-plot design

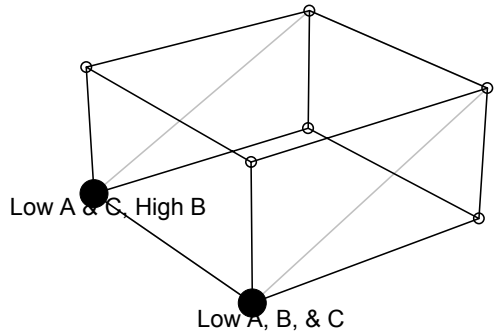
x	x	**A1**	**B1**	**C1**	**D1**	**B8**	**A8**	**D8**	**C8**	x	x
x	+	**A2**	**B2**	**C2**	**D2**	**B7**	**A7**	**D7**	**C7**	-	x
x	+	**A3**	**B3**	**C3**	**D3**	**B6**	**A6**	**D6**	**C6**	-	x
x	+	**A4**	**B4**	**C4**	**D4**	**B5**	**A5**	**D5**	**C5**	-	x
x	+	**A5**	**B5**	**C5**	**D5**	**B4**	**A4**	**D4**	**C4**	-	x
x	+	**A6**	**B6**	**C6**	**D6**	**B3**	**A3**	**D3**	**C3**	-	x
x	+	**A7**	**B7**	**C7**	**D7**	**B2**	**A2**	**D2**	**C2**	-	x
x	x	**A8**	**B8**	**D8**	**D8**	**B1**	**A1**	**D1**	**C1**	x	x

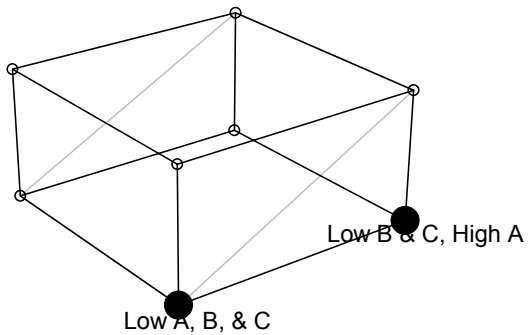
Uniformity trial #2



Designs: Factorial Experiments







-One Factor at a Time

A: Cell number

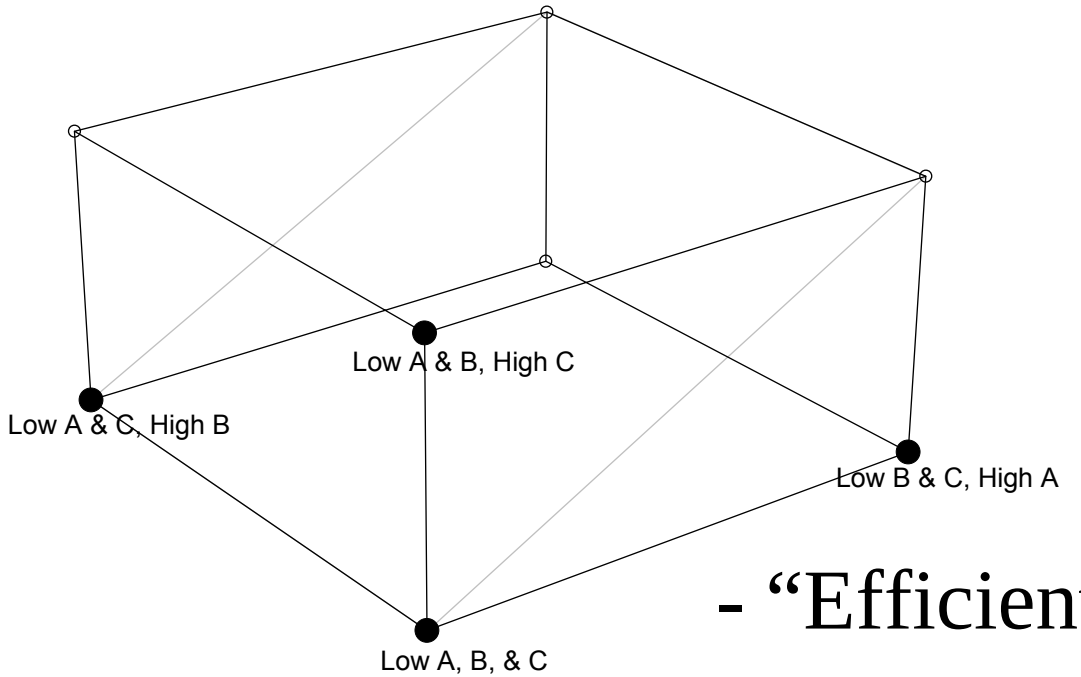
B: Virus number

C: Antibody amount

- 8 replicates of each pair

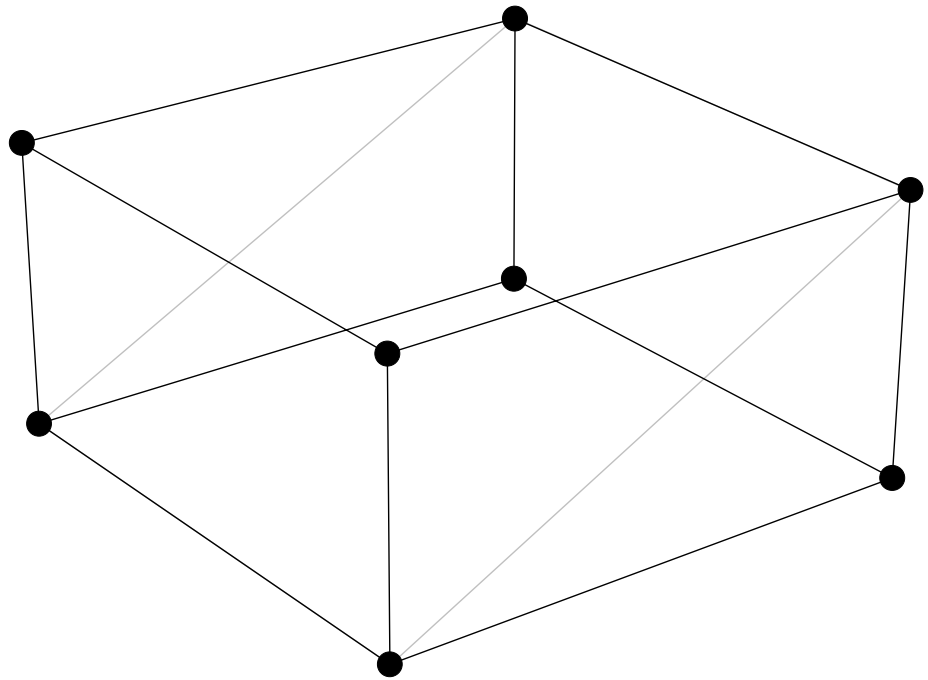
- Total of 48 experimental units

Designs: Factorial Experiments



- "Efficient" OFAT
- 8 reps at each point
- 24 EU total

Design: Factorial Experiments



- Full factorial (2^3) has 8 points

- Use 2 replicates of the full design

- Total of 16 experimental units

Designs: Factorial Experiments

Contrast	OFAT	“eff” OFAT	Full 2^3
A	8	8	8
B	8	8	8
C	8	8	8
A x B	0	0	8
A x C	0	0	8
B x C	0	0	8
A x B x C	0	0	8
Total EU	48	24	16

Designs: Factorial Experiments

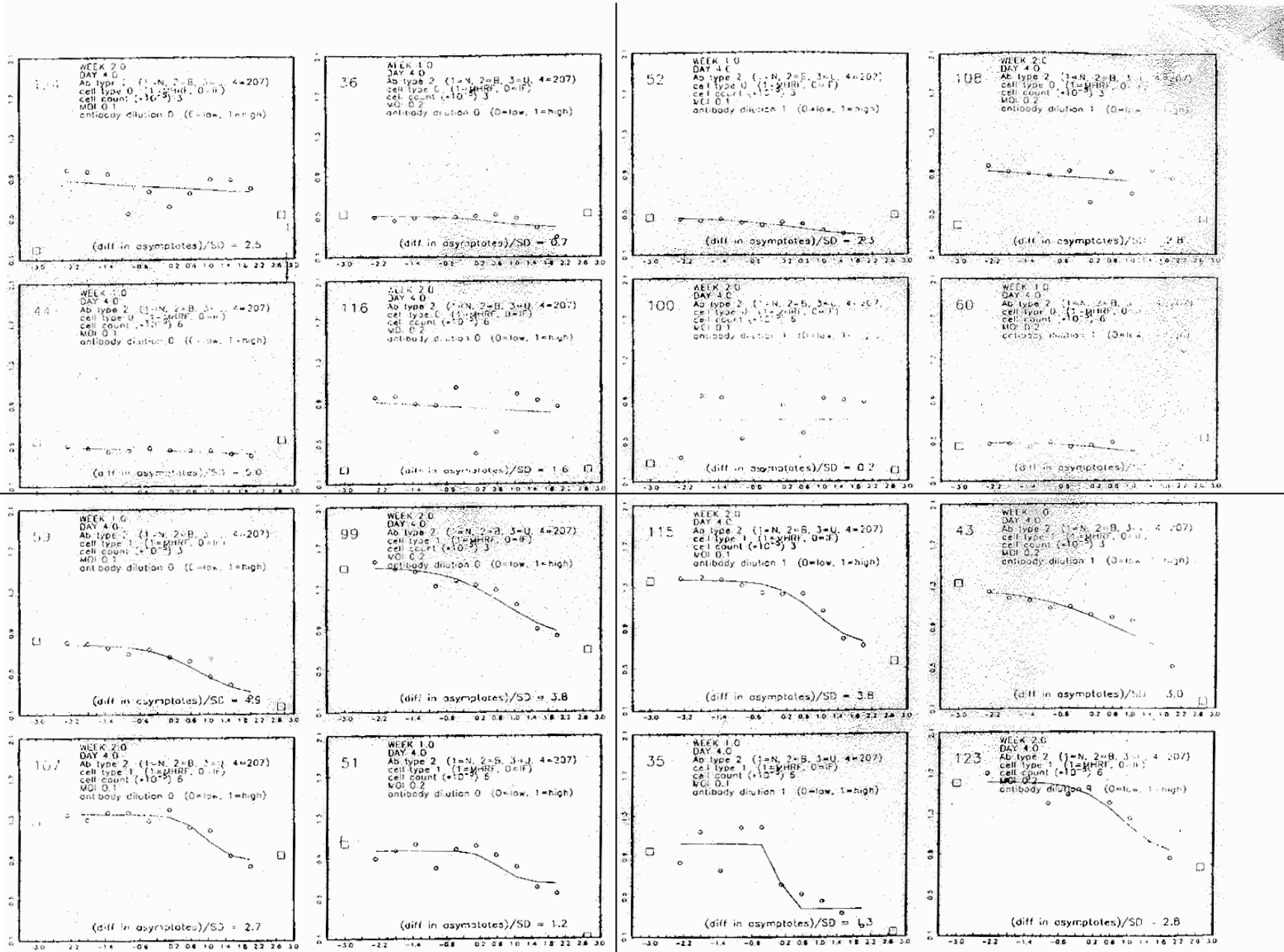
- Factorial experiments are vastly efficient
- Factorial experiments are the smart way to learn about interactions.
- The more factors you have, the more you gain from a factorial approach, BUT
- Full factorials get big fast with lots of factors
- Idea: don't run the full factorial
- With careful choices we can keep information we need and sacrifice the high order interactions
- See Montgomery, Douglas (2001) Design and Analysis of Experiments, 5th Edition, Wiley
- This is powerful, it fits well in fairly early development and in robustness testing

Assay Development

- Full factorial on a set of factors
  - 3 vs. 4 days
  - Antibody type (4 levels)
  - Cell type (IF vs MHRF)
  - Antibody amount (low and high)
  - Cell number (low & high)
  - Virus number (low & high)
- Goal: widen response range or reduce variation

$$t = \frac{\max - \min}{SD}$$

- One page for each cell type*antibody type



Assay Development Sequence

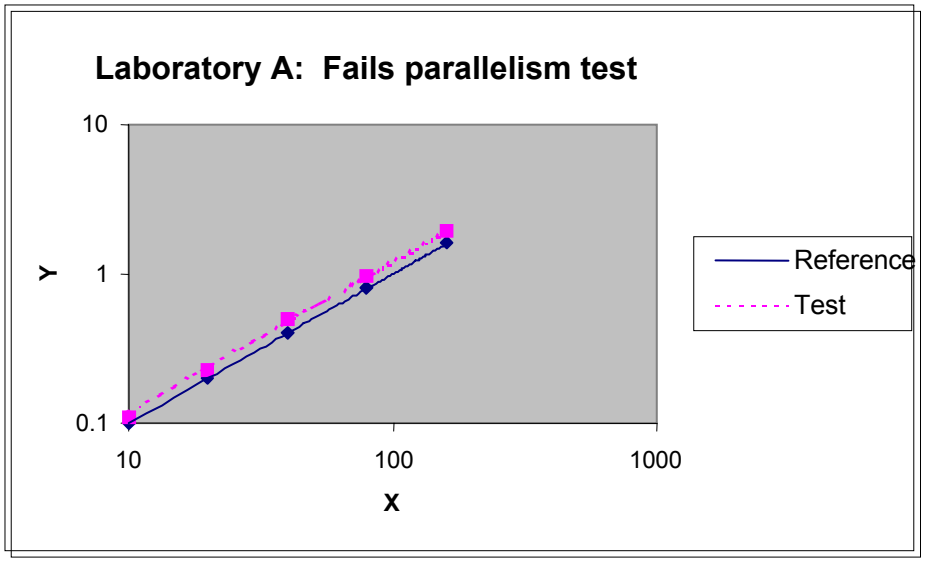
- If needed for initial range finding: OFAT
- Screen many factors with fractional factorial
- Full factorial with important factors in a design that yields variance information
- Response surface for optimization
- Fractional factorials for robustness
- Nested designs for validation

- Take a bioassay from 50% RSD to 13% RSD on potency in 6-12 weeks

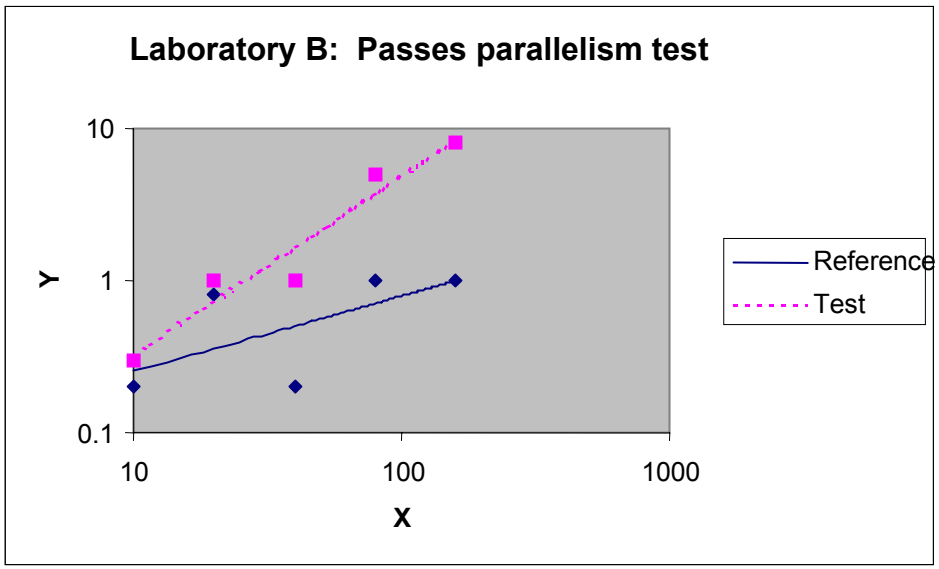
Why assess similarity in bioassay?

- Bioassay assumes
  - Test and reference contain the same active compound
  - Test and ref. differ only in conc. of active compound
- Without similarity relative potency has no meaning
- Similarity supports this key assumption
- Assessing slopes always appropriate, asymptote comparison essential for:
  - Stability
  - Comparison of assay methods
  - Comparison of references
  - Evaluation of changes in product

Ex: poor conclusions from parallelism tests



The run in Laboratory A fails the parallelism test because the low variability makes the test more sensitive

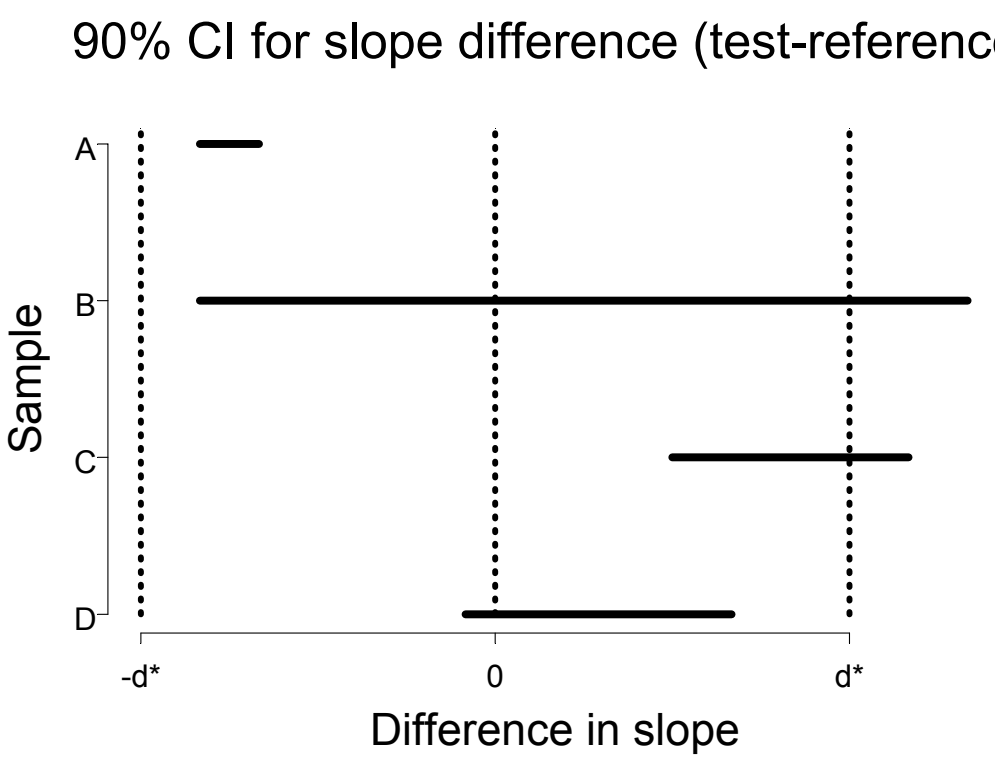


The run in Laboratory B passes the parallelism test because the high variability makes the test less sensitive

Parallelism tests have been set up incorrectly

- Statistical tests disprove a false H_0:
  - Set the false positive (or α) error rate (typically 5%)
  - Little or noisy data yields large false negative error rate – automatically conservative
- When we want to support H_0:
  - The false negative error rate is much more important than the false positive error rate
  - Not reasonable to expect to effectively control false negatives by fiddling with the false postive rate
  - Borrow technology from bioequivalence trials: use two one-sided tests (= confidence interval for difference with an indifference zone).

Using confidence intervals with an indifference zone to assess parallelism

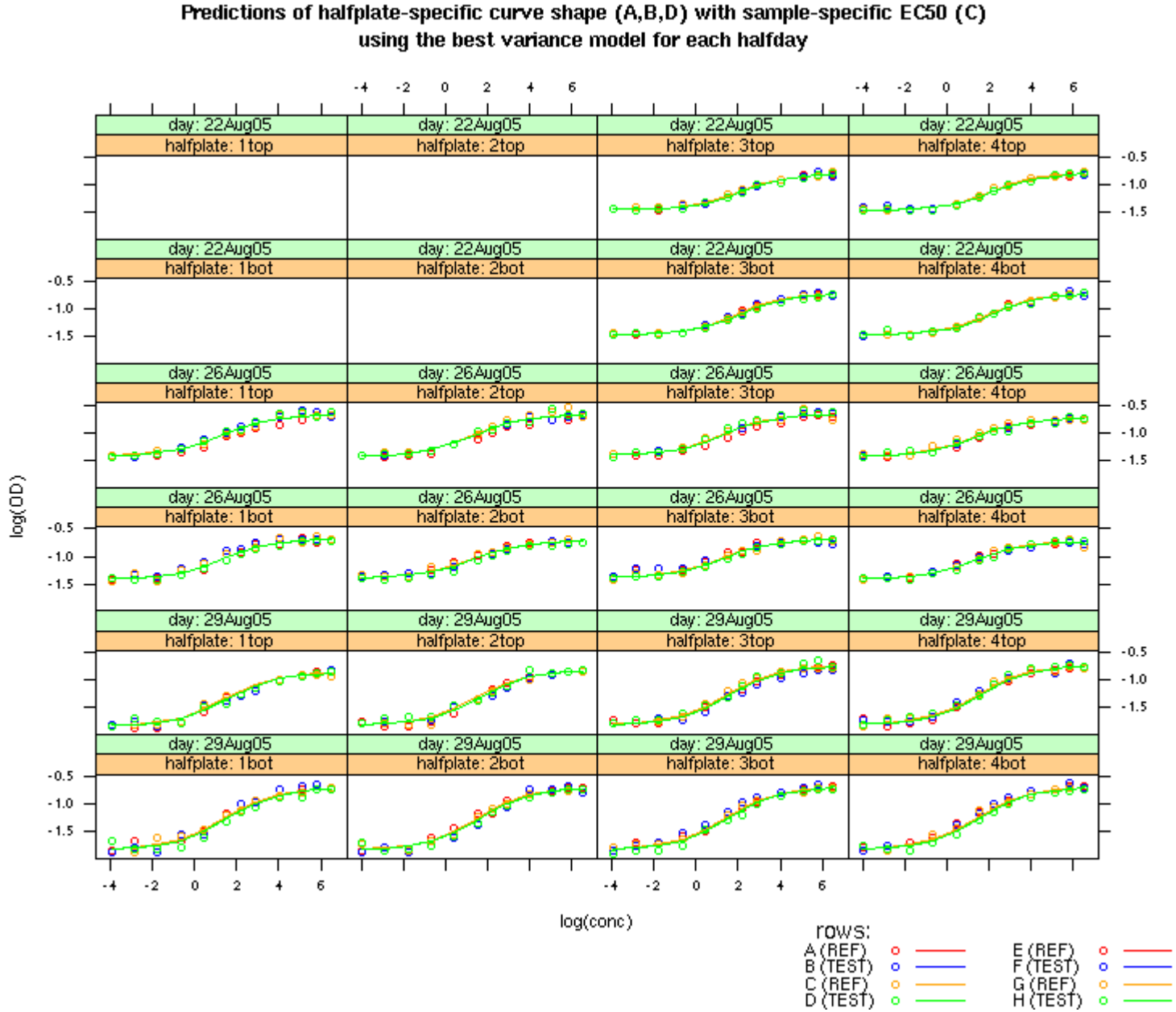


Does the sample PASS parallelism?

	Test	**CI**	**assay is:**
Sample A:	*No*	*Yes*	*precise*
Sample B:	*Yes*	*No*	*noisy*
Sample C:	*No*	*No*	
Sample D:	*Yes*	*Yes*	

Mixed models handle blocks well

- Most bioassays benefit from careful blocking
- Mixed models handle blocks well
  - This is particularly important when fitting linear models to bioassay data (which is fundamentally non-linear)
- For a good design (strip-plot) a good mixed model analysis has cut %RSD by more than half



Summary

- Good bioassay depends on
  - great biology
  - phenomenal animal care
  - careful use of statistics
  - Integration of all of the above
- Bioassay precision useful early
- Bioassay development
  - Factorial experiments can help a lot
  - Statistical training/coaching
  - Bioassay analysis software is often limiting